

## COMPARISON OF SPATIAL-TEMPORAL REGIONAL DEVELOPMENT AND SUSTAINABLE DEVELOPMENT STRATEGY IN LITHUANIA

Marija BURINSKIENĖ<sup>1</sup> and Vitalija RUDZKIENĖ<sup>2</sup>

<sup>1</sup> *Urban Engineering Department, Vilnius Gediminas Technical University, Saulėtekio al. 11, LT-10223 Vilnius, Lithuania*

*E-mail: marbur@ap.vtu.lt*

<sup>2</sup> *Department of Legal Informatics, Law University of Lithuania, Ateities 20, LT-2057 Vilnius, Lithuania*

*E-mail: vital@ltu.lt*

Received 8 June 2004; accepted 3 September 2004

**ABSTRACT.** Description of urban development could be explained like common endeavors of society to coordinate economical growth and social progress, to save usage of nonrenewable resources without threat of ecological balance. In the light of the attitudes of sustainable development researchers have raised the purpose to determine the current situation and possibilities of sustainable development of regions in Lithuania.

Sustainable urban and regional development is closely related with investment in tangible fixed assets as well as direct foreign investment. This paper describes an approach for recognition of investment information patterns and the main factors that have a distinct impact on investment. The proposed multivariate data analysis strategy allow to evaluate, forecast and establish the favorable conditions for investment and develop the strategies for cohesive development of regions.

In this article we present the application of the model using a database of economic and social development of regions in Lithuania. Our approach effectively recovers the structure of investment and enables forecasting. The capability to evaluate measurable features, identify the main features, and quantify their relationship with number of investments in space and time makes this approach applicable for decision support system and for creating range system for evaluation potency of the towns and regions.

**KEYWORDS:** Sustainable development; Multidimensional statistical methods; Significant factors; Regional planning

### 1. INTRODUCTION

The main aim of sustainable urban development is to mach up economical growth of the town progress, focusing on a more prudent consumption of natural resources, keeping the ecological balance as well as ensuring favorable living conditions for generations to come. UN Commission (Chief-G.H.Brundtland) emphasized that sustainable development should be based on not only the co-ordination of environmental and economical concerns, but it must

also ensure social justice inside the state and within international relation as well. However, the poverty is one of he most important obstacles while implementing sustainable development 16J. Sustainability is not a digital balance among all three aspects of the conception, their objectives and needs, although it is necessary to co-ordinate them and set prerequisites in order to implement the conception 18]. The most important features and requirements of towns and regions sustainable development were summarized in Agenda 21 [1].

The conception of sustainable development includes the way to reach two different and sometimes contradictory attitudes: "development-progress-grows" and "stability-security-environment" [8]. The problems of Lithuanian regions and towns go together with the sub-sequences that impact on social life of the town when some of the towns or regions degrade, while other economical processes lead the towns towards stagnation and they become unattractive for investments [14, 11]. Some depression regions emerge even regarding the territory of Lithuania. In addition, EU directives constantly highlight the importance of the regions and their equal development. During the last thirteen years uneven development was increasing. In general, the objective of sustainable development is to protect and improve quality of life. Preventive and needs management methods have a global influence on sustainability. The altered space of economy stimulated by globalization seems to support the political goals of spatial development outside core regions of Europe. A geographical theory of the economy of space presents the idea of guiding development by special policy measures [9]. The ESDP and VASAB 2010 are perspectives for special development in Baltic Sea region. The key policy issues of ESDP are the provision of:

- a polycentric spatial development,
- a new rural -urban relationship,
- equal access to infrastructure and knowledge,
- wise management of natural and cultural heritage.

With regard to the urban system, poly-centrism and co-operation are the key concepts of the ESDP. VASAB 2010 is a vision for the development of cities, infrastructure and nature conceptualized as "pearls", "strings" and "patches". Development and environmental sustainability as well as freedom and solidarity are promoted to be basic values of the development perspective. The balancing of these goals is a key issue of the VASAB 2010, which reveals international competitiveness and re-

gional cohesion as key measure. The cities are supposed to be responsible for economic development and social cohesion. "Improved links between urban centers and rural hinterlands shall allow all sub regions to participate in development process" (VASAB 2010, p. 10, section 3).

## **2. APPLICATION OF MULTIDIMENSIONAL STATISTICAL METHODS FOR DIRECT FOREIGN INVESTMENT ANALYSIS**

In order to describe the social-economic processes and phenomena, large sets of social-economic indicators are necessary. Most of these indicators take the form of time series in data warehouses. This causes some difficulties in connection with the establishment of interrelation structure of these indicators. In addition, many social and human-initiated events deal with incomplete or limited by nature information and a complex structure of interdependencies. That is why the use of statistical methods for the social-economic process analysis and decision-making is not only justified but also indispensable.

Multivariate statistical methods from the majority of possible probabilistic-statistical models enable us to make a grounded choice of a model that suits best to the initial statistical data that characterize the real behaviour of the set of objects under consideration and can estimate the present and future situation most exactly as well as to present substantiated and exact conclusions.

When describing the socio-economic situation, a great volume of initial data and indicators are used that characterize the development of a process. Therefore it is very important to select the most important ones and consider a small amount of indicators or their groups. Frequently the initial data are transformed in that way ensuring the minimal loss of information.

### 3. THE MODEL

Investment (like other human-initiated events) is random events in space and time. The process of dynamic structures recognition can be arranged in the following order:

(a) Observation objects of interest (towns, regions, districts, etc.) are selected, i.e., a sample  $O = (o_1, o_2, \dots, o_n)$ . The object of a data set is a unit of data whose features are to be investigated. The objects have respective features (or indicators)  $X = (x_1, x_2, \dots, x_n)$  that describe their attributes. These features are measured within particular time intervals (ranges, e.g., a year interval),

$$\Delta t = (\Delta t_1, \Delta t_2, \dots, \Delta t_k)$$

(b) Compose an  $(N \times n \times k)$  - dimensional matrix  $Q_{ijt}$  that consists of object features in the time intervals considered, where  $i$  is the object considered,  $j$  denotes measured features, and  $t$  is a time interval.

(c) Preparing data for a further analysis, we determine homogeneity of the objects observed by investigating their properties. To this end, the methods of cluster or variance analysis should be applied. Cluster analysis belongs to classification algorithms and solves an issue how to organize the observed data into meaningful structures. The general categories of the cluster analysis methods are joining or tree clustering, two-way joining or block clustering and k-means clustering.

If the clusters are clear heuristically, the methods of variance analysis are usually used. This classification problem can be solved in other ways, too: using heuristics or extreme way [3, 10]. Groups of objects  $N$  are defined by choosing a fixed time interval  $\Delta t$ , and soundness of the clusters formed is verified in other time intervals.

(d) When clusters of objects are formed, the structure of features characterizing the clusters is under determination. For this reasons the hierarchical factor analysis methods are selected for the problem solution [22]. The factor analysis is applied in order to reduce the number of variables and detect a structure in

the relationships between the variables. Generally, as a method for data reduction, the principal component analysis is often preferred, and the principal factor analysis is more frequently used in case when the goal of analysis is to detect structure.

(e) After verification of the data adequacy/suitability to the factor analysis, variables that are not suitable for the analysis are found and eliminated. The adequacy of data (variables) for the factor analysis can be verified with the help of Kaiser-Meyer-Olkin measure of sampling adequacy KMO [14, 17]:

$$KMO = \frac{\sum \sum_{i \neq j} r_{ij}}{\sum \sum_{i \neq j} r_{ij} + \sum \sum_{i \neq j} \tilde{r}_{ij}}, \quad (1)$$

where:  $r_{ij}$  is the correlation coefficient, and  $\tilde{r}_{ij}$  is the coefficient of partial correlation.

If the KMO value is low, then the indicators considered do not apply to the correlation analysis, since the correlation of these indicators cannot be explained by other indicators. For making the exploratory data analysis, it is recommended first to analyse principal components [19]. The components obtained through this analysis are not correlated and emerge in decreasing order of the amount of variance explained.

(f) To obtain a clear pattern of factor loadings, factor rotation strategies should be applied.

The fundamental theorem of factor analysis is invariant within rotations. The results of rotation, however, indicate "the simplest solution among a potentially infinite number of solutions that are equally compatible with the observed correlations" [18]. The simplest case of rotation is an orthogonal rotation. Typical orthogonal rotation strategies are *Varimax*, *Quartimax*, *Equamax*, and *Orthomax*. The *Varimax* rotation method is most commonly used orthogonal rotation procedure. The overriding criterion of simple structure is that each factor should have a few high loadings with the rest being zero or close to zero [19]. After clearing the patterns of factors, the influence of in-

dividual indicators  $x_1, x_2, \dots, x_n$  is evaluated and the factor interpretation is performed.

(g) The interdependence and significance of variables (indicators) composing the factors is evaluated through developing a multivariate regression equation for time intervals  $At$ . A multiple regression analysis determines the relationship between several independent variables and a dependent variable. The regression function can be estimated, using the least squares estimation or any other loss function (non-linear estimation). After the regression equation has been estimated, the prediction can be computed for a set of independent variables.

(h) Thereafter the factors are defined and the number of variables is reduced the observed objects should be organized into meaningful structures. For this purpose the cluster analysis algorithms have to be applied. Cluster analysis solves an issue how to organize the observed data into meaningful structures that is, to develop taxonomies. The general categories of the cluster analysis methods are such as joining or tree clustering, two-way joining or block clustering and k-means clustering.

#### 4. MODEL EVALUATION

In the selection process of observation objects of interest the set of 13 social- economic indicators were collected for the research from 12 Lithuanian towns and 43 regions during time intervals of the period from 1996 until 2001 [5]. We consider the matrix denotes as  $X[n*N]$ . The matrix elements  $x_{ij}$  illustrate the value of the  $j^{\text{th}}$  indicator at the  $i^{\text{th}}$  research object and have particular values and semantics:

- $x_{i,1}$  - registered crimes;
- $x_{i,2}$  - average annual number of employed;
- $x_{i,3}$  - unemployment rate;
- $x_{i,4}$  - natural increase;
- $x_{i,5}$  - migration;
- $x_{i,6}$  - average monthly gross earnings;
- $x_{i,7}$  - sales of industrial production;
- $x_{i,8}$  - average real estate price;
- $x_{i,9}$  - dwelling acquisition;
- $x_{i,10}$  - investments in the construction of

residential houses;

$x_{i,11}$  - investment in tangible fixed assets;

$x_{i,12}$  - direct foreign investment;

$x_{i,13}$  - turnover of catering, where:

$i=1,2,\dots,N$ .

The objective of the research was to explore, estimate, and apply the use of multivariate statistical models in the analysis and prediction of the state situation and tendencies for even distribution of the quality of life in Lithuanian towns and regions, paying particular attention to the determination of the possibilities of sustainable regional development in Lithuania. In order to estimate the situation and make decisions it is expedient to evaluate and select the main factors that have an influence upon direct foreign investment in Lithuanian towns and regions. For this purpose, the factor and component analysis are most frequently used. These methods make it possible to evaluate the multidimensionality of essential data and explain concisely and simply enough the multivariate structures. They reveal really existing, but directly imperceptible regularities by means of factors or principal components.

The aim of factor analysis is to explain the outcome of  $p$  variables in the data matrix  $X$  using fewer variables, the so-called factors. These factors are interpreted as latent (unobserved) common characteristics of the observed  $x \in R^n$ . In the factor analysis every observed  $x = (x_1, \dots, x_n)$  can be written as

$$x_j = \sum_{l=1}^k a_{jl} f_l + \epsilon_j, \quad j=1, \dots, n; k \leq n, \quad (2)$$

where:  $f_l$  for  $l=1, \dots, k$  denotes the factors,  $e$  is the residual of  $x$  on the factors. Given the assumption that the residuals are uncorrected across the observed variables, the correlations among the observed variables are accounted for the factors.

When using the factor analysis we face a number of problems. Most often they emerge in data suitability testing for the factor analysis, when choosing and interpreting factors. According to the logical sequence of problems solved by the factor analysis, the arising prob-

lems can be arranged in the following order: the first problem is robustness, the second one - community, the third - factors, the fourth -rotation, the fifth - estimation of factor values, and the-sixth problem - dynamic models [10]

When the five problems listed above are solved,, we regard the problem of factor analysis as solved.

The sixth problem arises when we use the factor analysis for data of more than a year, comparing and interpreting the obtained results in time. This problem is related with the principles of building dynamic models. Dynamic models render on opportunity to find out indications, the influence of which may decrease in future, or, on the opposite, would increase. Dynamic models enable us to compare or predict the results of factor analysis in a certain segment of time.

Several important issues are considered preparing data for the factor analysis. Firstly, it is important to know which variables should be included for analyse. Secondly, how many variables should be included. A factor cannot be defined using a single observed variable. There should be a minimum of three observed variables for each factor expected to emerge [13]. Thirdly, the number of observations should be sufficient to provide reliable estimations of the correlations between the variables. Correlation coefficients tend to be unstable and greatly influenced by the presence of outliers if the

sample size is not large. There is a general opinion that it is unwise to conduct the factor analysis on a sample fewer 50 observations. Another recommendation is that the sample size should be twice as large as the amount of variables, that is, there should be at least twice as many cases (N) as variables (n) and the cases should be heterogeneous with respect to the measures studied [19].

While considering the Lithuanian social-economic indices of 1996-2001, the sample of objects studied has naturally to be divided into two groups: the first group consists of the largest cities and resort towns, and the second one - of regions. In order to form the groups, we can use cluster analysis methods, however, in this particular case, group bounds are clear. Substantiation of the division is verified by the hypothesis  $H_0$  stating that the average number of direct foreign investment in towns and regions is equal. This hypothesis is verified by the criterion:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_1^2/N_1 + S_2^2/N_2}}, \quad (3)$$

where:  $x$  is the estimate of mean, and  $S$  is the standard deviation.

Arithmetic means of the direct foreign investment calculated, values of the criterion  $t$ , degrees of freedom, and the observed significance level  $p$  are presented in Table 1.

According to the obtained results shown in Table 1, the significance level observed in the

Table 1. Verification results of the hypothesis that the number of direct foreign investment in towns and regions is the same

Year	Average number of direct foreign investment in towns	Average number of direct foreign investment in regions	T-value	Df	p
1996	451.4	155.5	2.12	44	0.0397
1997	876.7	234.1	3.14	49	0.0028
1998	1384.5	288.0	4.21	48	0.0001
1999	1927.1	415.8	3.97	47	0.0002
2000	2421.1	438.8	4.09	48	0.0002
2001	2412.1	517.5	3.45	50	0.0011

years under investigation is lower than 5%. Therefore we have to reject the hypothesis  $H_0$  and to consider the direct foreign investment in towns and regions as separate.

After evaluating the influence of each variable with the help of KMO measure, we eliminated four variables from the list of indices considered, namely: average annual number of employed, unemployment rate, natural increase, and investments in the construction of residential houses. KMO measure of the rest variables  $KMO=0.68$ , thus, we concluded by stating that the data are adequate to the factorial analysis. In order to make the exploratory data analysis, it is recommended first to analyze principal components (19). The components obtained via this analysis are not correlated and emerge in decreasing order of the amount of variance explained.

The number of factors to be extracted can be determined in a scree plot (Figure 1).

The first large eigenvalue (2.9) and the second much smaller eigenvalue (1.52) suggest the presence of a dominant global factor. The most widely used criterion for finding number of factors is the Kaiser criterion, which recommends to retain only the factors whose eigenvalues

are greater than 1 [14]. The scree plot (Figure 2) suggests a maximum of four factors, too. These four factors account for 64.3% of the whole variance.

After evaluating the number of factors to be extracted, the next logical step is to determine the method of rotation. The overriding criterion of simple structure is that each factor should have a few high loadings with the rest being zero or close to zero [15]. Applying this criterion the Biquartimax method was selected as providing the simplest structure solution.

When the rotation method is applied, one part of the output from the factor analysis is a matrix of factor loadings (Table 2). A factor loadings or factor structure matrix is a matrix of correlations between the original variables and their factors.

#### 4.1. Interpretation of factors

The meaning of the rotated factors is inferred from the variables significantly loaded on their factors. A decision has to be made regarding what constitute a significant loading. The simplest criterion is that factors loadings

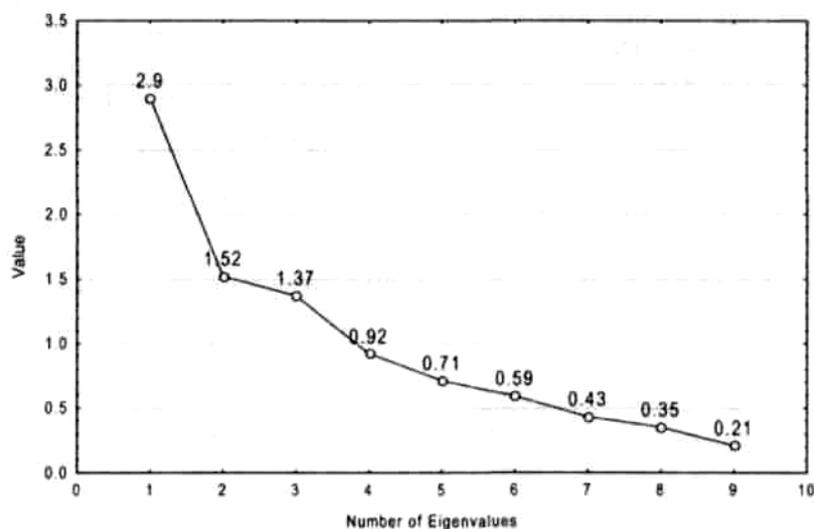


Figure 1. The rate of change in the magnitude of eigenvalues for the factors

**Table 2.** Factor Loadings (clusters of loadings are marked)

Factor Loadings (Biquartimax normalized)	Factor 1	Factor 2	Factor 3
Direct foreign investment	<b>0.66</b>	0.26	0.02
Migration	-0.09	<b>0.72</b>	-0.01
Average monthly gross earnings	0.53	0.03	<b>-0.67</b>
Sales of industrial production	<b>0.84</b>	-0.28	0.16
Average real estate price	0.13	<b>0.80</b>	0.15
Dwelling acquisition	-0.24	-0.17	<b>-0.85</b>
Investment in tangible fixed assets	<b>0.77</b>	0.22	-0.29
Turnover of catering	<b>0.81</b>	-0.24	0.19
Registered crimes	0.33	0.25	-0.23
Expl. Var	2.88	1.51	1.39

greater than 0.30 in absolute value are considered to be significant. As the sample size increases, the criterion may need to be adjusted a bit downward. When the number of factors increases it may be adjusted upward. In general, the larger the absolute size of the factor loading for a variable, the more important the variable is in interpreting the factor.

As we can see from results in Table 2, the most significant variables for the first factor are as follows:

- Direct foreign investment.
- Sales of industrial production.
- Investment in tangible fixed assets.
- Turnover of catering.

It should be stated that the first factor reflects the growth of economy and the improving life conditions of many people. The greatest impact on this factor is made by the variables such as *sales of industrial production* ( $L=0.84$ ), *turnover of catering* ( $L=0.81$ ), that reflect the increasing retail trade, *investment in tangible fixed assets* ( $L=0.77$ ), and *direct foreign investment* ( $L=0.773$ ) which indicates the increasing influence of foreign investment.

The second factor reflects people adaptability to changing circumstances and it consists of two variables such as *migration* ( $L=0.72$ ) and *the average real estate price* ( $L=0.80$ ).

The third factor has two variables, *average monthly gross earnings* ( $L=-0.6$ ) and *dwelling acquisition* ( $L=-0.8$ ). This factor shows improving economic situation in Lithuania and relation between average monthly gross earnings and dwelling acquisition.

#### 4.2. Factors interdependence

After selecting the principle variables that describe the economic development of the state, according to the number of these and the direct foreign investment, the system determine their mutual relationship and its change in time. For this we will make use of a linear equation of multiple regressions:

$$y_i = \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i, i = 1, \dots, N, \quad (4)$$

where:  $\beta$  is unknown parameters;  $\varepsilon$  is a random variable with zero arithmetic average mean and variance  $\delta^2$ .

For the first factor  $k=3$ ,  $y$  denotes *direct foreign investment*,  $x_1$  is *sales of industrial production*,  $x_2$  is *investment in tangible fixed assets*,  $x_3$  denotes *turnover of catering*.

In order to verify the significance of the regression equation, the variance analysis criterion  $F$  is used. The F-value and the resulting

p-value reflect the relationship between the dependent variable and independent variables. Having calculated the value of this criterion for all the data of years analyzed, we obtain  $F(2,234)=37.2$  and  $p<0.0001$ . The value of multivariate correlation coefficient is  $R=0.51$ , and of the determination coefficient  $R^2=0.25$ .

The significance of the regression coefficients, i.e. the hypothesis  $H_0: \beta=0$  is verified by the criterion  $t_j$ :

$$t_j = \frac{b_j}{[\hat{\sigma}^2 C_{jj}]^{1/2}}, \quad (5)$$

where:  $C_{jj}$  is the diagonal element of  $(X^T X)^{-1}$  corresponding to  $b_j$ . In the case of the hypotheses  $H_0: \beta=0$  being correct,  $t_j$  is distributed by Student law with an  $n-k-1$  degree of freedom.

For the variable *sales of industrial production* at the 5 percent level of significance the null hypothesis should be accepted. The variables  $x_2$  is *investment in tangible fixed assets*,  $x_3$  denotes *turnover of catering* are statistically significant at the 5 percent level of significance.

In order to learn which of the selected indicators contributes most to the prediction of direct foreign investment, we need to examine the standardized regression coefficients. These coefficients are obtained by standardizing the variables to a mean of 0 and the standard deviation of 1.

Thus, the magnitude of these coefficients allows a comparison of a relative contribution

of each independent variable in the prediction of dependent variables. The coefficients are obtained separately for 1996-2001 (Figure 2).

According to the results obtained through the regression analysis, we can state that the investment in tangible fixed assets *acquires more weight, which has the greatest importance in prediction of investment*.

Analyzing the second and third factors in the same way, we find that the *variables average real estate price and average monthly gross earnings and dwelling acquisition* have statistically significant relation with direct foreign investment.

Examining the standardized regression coefficients for the third factor for 1996-2001 we can confirm the stable relation (Figure 2). The results demonstrate that the relation between the variables *average monthly gross earnings, dwelling acquisition, and direct foreign investment* remains significant almost during the whole period.

#### 4.3. Identification of clusters members

The purpose of cluster analysis in this research is to find regions with similar characteristics. Evidence for the new faith in the economy of space can be found in theories of creating regional competitiveness by localized learning, development of governance leadership and by development of clusters. Theories of clustering are *par excellence* theories of economy

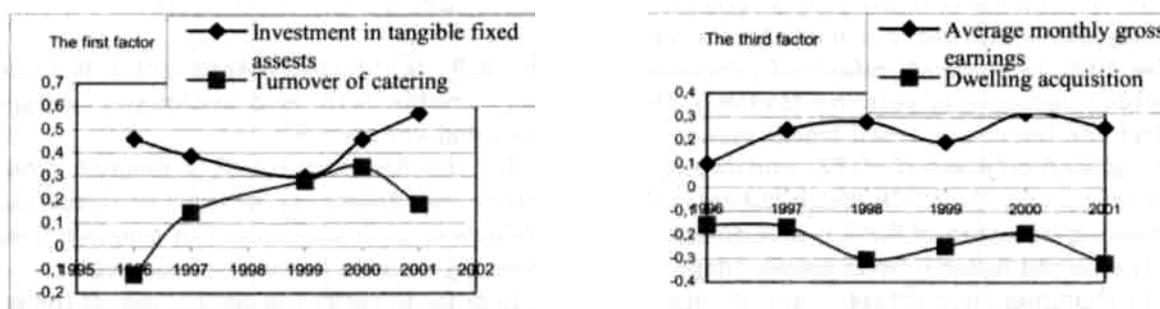


Figure 2. Standardized regression coefficients in 1996 - 2001 confirm the stable influence of the variables of the first and third factor to the direct foreign investment

of space since they rely on the assumption that geographical proximity between related production units creates added value and local competitiveness [9].

Cluster analysis attempts to identify relatively homogeneous groups of cases (or variables) based on selected characteristics, using an algorithm that starts with each case (or variable) in a separate cluster and combines clusters until only one is left [12]. In the case when a large number of variables are used for cluster analysis it is recommended to reduce the number of variables before starting cluster analysis procedure. The factor analysis often is used as one of methods for reducing the number of variables [3]. We will use factor scores estimates for the regions' classification.

There are two general classes of methods for estimating factor scores. The first class of methods yields approximately standardized factor score estimate with different properties. Regression approach produces factor score estimates that maximize determinacy [2].

$$F = \Phi \cdot \Lambda^T \Sigma^{-1} X, \quad (6)$$

where:  $F$  is estimated common factors;  $\Phi$  is the covariance matrix of the common factors;  $\Lambda$  is the matrix of loadings;  $\Sigma$  is the model-implied covariance matrix of the measured loadings. Matrices are based on estimated parameters.

Another methods yields factor score estimates that are perfectly orthogonal (uncorrected) [20]. Each of the refined methods is imperfect. Regression estimates will be correlated even when the factors are orthogonal, and orthogonal estimates will not maximize determinacy.

Choosing among the different factor scoring procedures is thus not a straightforward affair, and the final choice must be dictated by a survey of the extant studies on the differences among the various estimation methods and by the context of particular research program.

Having computed the regression estimates of factor scores, the data were partitioned by

separate years and performed agglomerative hierarchical cluster analysis. Creating clusters with the help of this method, each case starts out as a cluster. At every step, clusters are combined until all cases are members of a single cluster.

Squared Euclidean distance was chosen as the measure of classification. This distance is computed as

Agglomerative hierarchical clustering helped to determine the number of clusters. Applying this procedure were determined that the optimal number of clusters is five clusters. Dividing the regions into five clusters we obtain such cluster membership (Table 3).

The location of the regions in Table 3 represents the social-economic development of the region or groups of regions as well as the capability to attract foreign investment. The most developed and attractive are regions in the top of the columns and less attractive are in the bottom. The results of clustering analysis show that the most developed regions are located around biggest cities and the main plants, important for the whole of Lithuanian economy. Comparing the groups (clusters) created in 1996, 1999, and 2001, we can notice that the situation is almost stable. It shows that direct foreign investments tend to concentrate in the regions near separate towns and determine their perspective development.

## **5. INTERPRETATION OF CLUSTER ANALYSIS RESULTS**

Since 1960, Lithuanian regional planning followed the concept of the universal settlement system, based on Christaller's theory of central places. In 1970, the paradigm stressed the role of urban centers and their modernizing effect on the periphery. The tools of regional planning have changed completely after the regaining independence [16]. The Act on Territorial Planning (1995) defines the levels of ter-

Table 3. Cluster membership of regions for factor scores in 1996-2001\*

1996	1997	1998	1999	2000	2001
Trakai (5)	Mažeikiai (5)	Mažeikiai (5)	Mažeikiai (5)	Mažeikiai (5)	Vilnius (5)
Mažeikiai (4)	Klaipėda (4)	Kaunas (4)	Kėdainiai (4)	Panevėžys (4)	Mažeikiai (4)
Kaunas (3)	Trakai (4)	Klaipėda (4)	Klaipėda (4)	Kaunas (3)	Klaipėda (3)
Klaipėda (3)	Ignalina (3)	Kretinga (4)	Utena (4)	Kėdainiai (3)	Ignalina (2)
Vilnius (3)	Alytus (2)	Šiauliai (4)	Kaunas (3)	Klaipėda (3)	
Ignalina (2)	Kaunas (2)	Trakai (4)	Kretinga (3)	Kretinga (3)	
	Kretinga (2)	Vilnius (4)	Trakai (3)	Trakai (3)	
	Vilnius (2)	Ignalina (3)	Vilnius (3)	Utena (3)	
		Alytus (2)	Ignalina (2)	Vilnius (3)	
		Anykščiai (2)		Ignalina (2)	
		Panevėžys (2)			

\* In brackets cluster numbers are marked. Regions that are out of listing depend to the first cluster.

territorial planning: the nation, the county and the municipality. All levels can elaborate comprehensive plans and special plans for subsystems, such as water supply or transportation development. The detailed planning is carried out at the municipality level. The Comprehensive Plan for the Lithuanian territory was approved by the Lithuanian Parliament in 2002. This Plan defined the national guidelines for spatial planning and support the implementation of regional policy [7]. At this moment it is the main document for physical planning and also created preconditions for sustainable development in the whole territory of Lithuania. In 2002, Strategic Plans for economy sectors development were finalized [21]. The connection of these strategic documents has created background to implement sustainable development in Lithuanian regions.

Due to post-war regional planning, Lithuanian urban system was well balanced, if a balance is understood as a graduated city ranging and an equal dispersion of the centers in the territory. But now in Lithuania the wealth is becoming increasingly concentrated in the capital city and the regions with main plants. This has lifted the position of Vilnius when compared with the situation before the

regaining of independence. The current goal -the ESDP of a polycentric urban system has similarities with the post-war regional planning in Lithuania. The cluster analysis used to evaluate the regional development in Lithuania allows showing changes of the last years (Figures 3, 4, 5).

The main principles for the regional policy were presented as a result of a cross-sectorial approach. There was a comprehensive result of this plan graphically expressed in the following main schemes [7]:

- Macro-regional situation of Lithuania,
- Spatial concept of the territory: main territorial structures and principal model.
- Functional priorities of the territory,
- Development of the technical infrastructure. The comparison of this scheme (Figure 6)

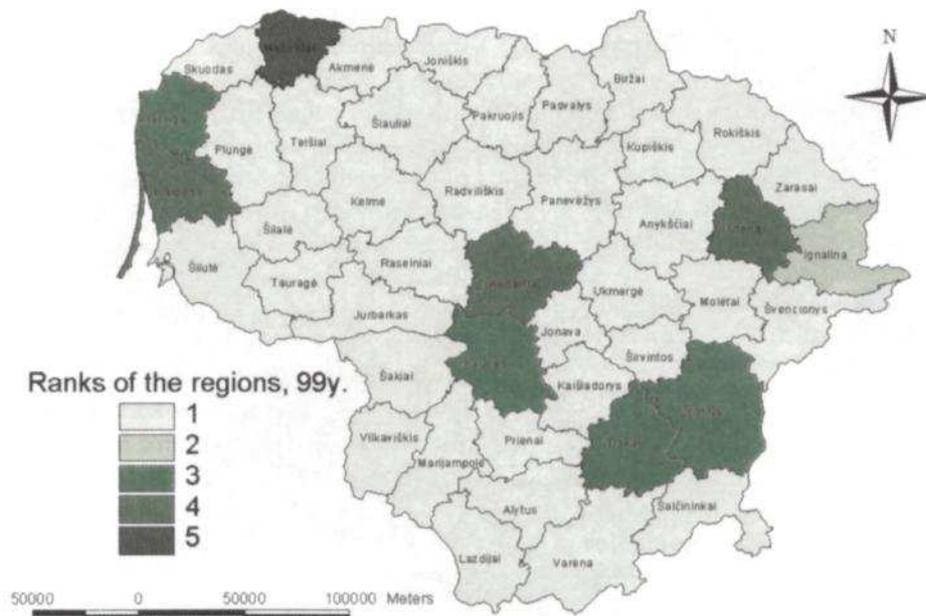
with ranks of the Lithuanian regions, identified by the methods of factor and clusters analysis according to followed years allows evaluating sequences of Lithuanian regional development.

For example, the region of Vilnius since 1996 till 2001 had better position in respect of other regions of Lithuania, and it reached the highest cluster in 2001. On the other hand, during this period Kaunas became more attractive and

*Comparison of Spatial-Temporal Regional Development and Sustainable Development Strategy in Lithuania*



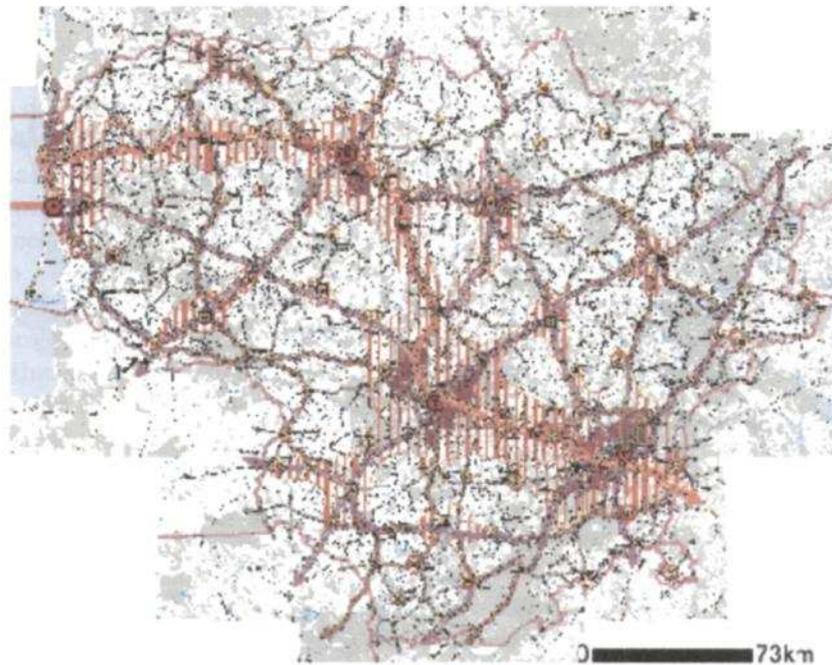
**Figure 3.** Ranks of the social-economic development and the capability to attract foreign investment of the Lithuanian regions identified by the methods of clusters analysis in 1996



**Figure 4.** Ranks of the Lithuanian regions identified by the methods of clusters analysis in 1999



**Figure 5.** Ranks of the Lithuanian regions in 2001



**Figure 6.** The main settlements framework according to Comprehensive Plan of the Territory of the Republic of Lithuania

in 2001 the region of Kaunas came up to the most of Lithuanian country regions. Total amount of direct foreign investment for the whole country increases from year to year, however, in last period foreign investments are centered on four main regions: Vilnius - the Capital, Klaipėda - Sea port, Mažeikiai - Oil production plant and Ignalina Nuclear Power Plant. Other regions were indicated in cluster membership table with higher numbers than 2, shows higher ability to attract and reclaim investments than other regions.

## 6. CONCLUSIONS

When analyzing the socio-economic situation, we have to use many interrelated initial data and indicators that characterize the development of the process. In employing multivariate statistical methods from many possible probabilistic - statistical models there was obtained the model such that describes the real behavior of the explored set of objects best and that provides substantiated and exact conclusions.

In order to evaluate the situation and make decisions the described knowledge discovery process enables the evaluation of main factors and selection the influencing space of the direct foreign investment in regions of Lithuania. During last period the territorial distribution of the investment is not even. To this purposes the integration of factorial and component analysis methods have been used. These models allowed the estimation of essential data multidimensionality and a concise and simplified explanation of multivariate structures of data. The regularities were revealed by means of factors, principal components and multivariate regression analysis existing in reality but directly imperceptible.

The analysis of social-economic data during the period of 1996-2001 shows that higher development was reached in the regions located close to the biggest cities and main plants, important for the whole of Lithuanian economy. The development of Vilnius, Klaipėda, and

Kaunas, level of investment in tangible fixed assets of these cities had influence to their region development. Ignalina Nuclear Power Plant, Mažeikiai Oil production plant had reflection to Mažeikiai and Ignalina region development, increase quality of life in these areas. Other regions were indicated in cluster membership table with higher numbers than 2, shows higher ability to attract and reclaim investments than other regions.

A comparison the Comprehensive Plan for the Lithuanian territory with ranks of the Lithuanian regions, identified by the methods of factor and clusters analysis from 1996 till 2001 allows evaluating sequences of development of Lithuanian regions. The statistical analysis shows that necessary to change investments policy and creating legal and economic directives for investments regulation, without these measures investment will be concentrated in regions near largest cities and whole Lithuania important plants, herewith increasing the gap between cities and peripheral towns as well as these regions, that goes in opposite towards sustainable development.

## REFERENCES

- [1] *An Agenda 21 for the Baltic Sea region* - Baltic 21 series, Nol/98. 34 p.
- [2] K. A. Bollen, *Structural Equation with Latent Variables*, Wiley-Interscience, 1989.
- [3] A. Buhl, P. Zofel, *SPSS Version 10*, Einführung in die moderne Dateanalyse uner Widows, Addison-Wesley, 2000.
- [4] M. Burinskienė, D. Dzemydienė, V. Rudzkienė, An Approach for Recognition of Significant Factors for Sustainable Development Strategies, in *Proceedings of the International Conference on Modelling and Simulation Business Systems*, Vilnius, 2003, p. 90-96.
- [5] *Counties of Lithuania: Economic and Social development*, Department of Statistics of Lithuania, Vilnius, 2002.
- [6] R. Čiegis, *Sustainable development and environment*, Economical view, Vilnius, 2002, 692 p. (In Lithuanian).

- [7] Comprehensive Plan of the Territory of the Republic of Lithuania. Approved by Lithuanian Parliament at 2002 10 29, No. 9-1154, Valstybes žinios 2002 U 15, No. 110-4852. (in Lithuanian).
- [8] V. I. Danilov-Daniljan, K.S. Losev, *Ecological Challenge and sustainable development*, Moscow: Progress-Tradition, 2000, 416 p. (In Russian).
- [9] M. W. Danson, Debates and Surveys, *Regional Studies*, 34(6), 2000, p. 571-580.
- [10] A. M. Dubrov, V. S. Mchitarian, L. I. Troshin, *Multiple statistical methods*, Moscow, Finance and Statistic, 1998. (In Russian).
- [11] D. Dzemydienė, V. Rudzkiene. Multiple Regression Analysis of Crime Pattern Warehouse for Decision Support, *Lecture Note in Computer Science*, Vol. 2453, Database and Expert systems Applications, Springer, 2002, p. 249-258.
- [12] B. S. Everit, Cluster Analysis, New York: Wiley, 1993.
- [13] L. Guttman, Image Theory for the Structure of Quatitative Variables, *Pyrometrical*, 1953, 18, p. 277-296.
- [14] A. R. Hakstian, W.D. Rogers, R.B. Cattell, The behavior of numbers of factors rules with simulated data, *Multivariate Behavioral Research*, 17, 1982, p. 193-219.
- [15] H. F. Kaiser, The Varimax Criterion for Analytic Rotation in Factor Analysis, *Pyrometrical*, 1958, 23, p. 187-200.
- [16] Sc-Th. Kaisa, B. Christer, ESDP and Spatial Planning and development in the Baltic Countries, Stockholm, Nordregio, 1999, (Nordregio Report 1999:2) 89 p.
- [17] H. F. Kaiser, The Application of Electronic Computers to Factor Analysis, *Educational and Psychological Measurement*, 20, 1960, p. 141-151.
- [18] J. O. Kim, C. V Mueller, *Introduction to Factor Analysis*, Beverly Hills: Sage Publications, 1978.
- [19] P. Kline, *An Easy guide to Factor Analysis*, London: Routledge, 1994.
- [20] W.P Krijnen, T Wansbeek, J. M. F. Ten Berge, Best Linear Predictors for Factor Scores, *Communications in Statistics: Theory and Methods*, 25, 1996, p. 3013-3025.
- [21] Long-tei-m Economic Development Strategy of Lithuania until 2015, Vilnius, 2003, 192 p. (in Lithuanian), [www.ukmin.lt/catalogs/33/strategijos/galutine](http://www.ukmin.lt/catalogs/33/strategijos/galutine)
- [22] R. J. Wherry, Contributions to correlation analysis, *New York: Academic Press, 1984*.